**Contents:**

1. Python 3.7.1
2. Visual C++ build tools
3. Visual Studio Code +
   1. Magic Python
4. Dbeaver
5. SQLAlchmey 1.2.14
6. Psycopg2 (version 2.7.6.1)
7. future 0.17.1
8. Pyinstaller ==3.5
   1. Pywin32.ctypes == 0.2.0
   2. Pywin32 ==224
   3. Macholib ==1.11
   4. Altgraph ==0.16.1
   5. Pefile =2017.8.1
9. Rundeck required software’s
   1. Java 11.1
   2. Nssm.exe 2.24
   3. rundeck-3.0.9-20181127

10. Oracle Remote Database Connection through Python

Pre- requirements

10.1 instantclient-basic-windows.x64-12.1.0.2.0

10.2 instantclient-odbc-windows.x64-12.1.0.2.0

10.3 **VS 2010**

11. instantclient-sqlplus-windows.x64-12.1.0.2.0

12. Pandas

12.1 Six

12.2 Python –dateutil

12.3 Pytz

12.4 Psutil

12. Numpy

13. SQL server jar files

13.1 jtds-1.3.1.jar

13.2 ntlmauth.dll

13.3 x64\SSO\ ntlmauth.dll

13.4 x86\SSO\ ntlmauth.dll

14. Postgres

14.1 postgis-jdbc-2.2.1

14.2 postgis-jdbc-jtsparser-2.2.1

14.3 postgresql-42.2.5.jre7

15. MySQL

15.1 mysql-connector-java-5.1.47\_1.jar (CLIENT)

15.2 mysql-connector-python-8.0.15-py3.7-windows-x86-64bit(PYTHON)

**1. Introduction**

**1.1. Purpose**

This document describes all the installation process for running ETL scripts. Software starting from scratch. This document covers all aspect of its installation including all various software’s needed, and purpose of the packages.

**1.2. How to Use This Document**

When installing the ETL Software’s from scratch, readers are encouraged to follow all the various software that needs to be installed before going to run the ETL scripts.

**2. Python Software Installation**

**2.1 Purpose of Python**

Python is a general purpose and high level programming language. You can use Python for developing desktop GUI applications, websites and web applications. Also, Python, as a high level programming language, allows you to focus on core functionality of the application by taking care of common programming tasks.

**2.2 Python installation steps**

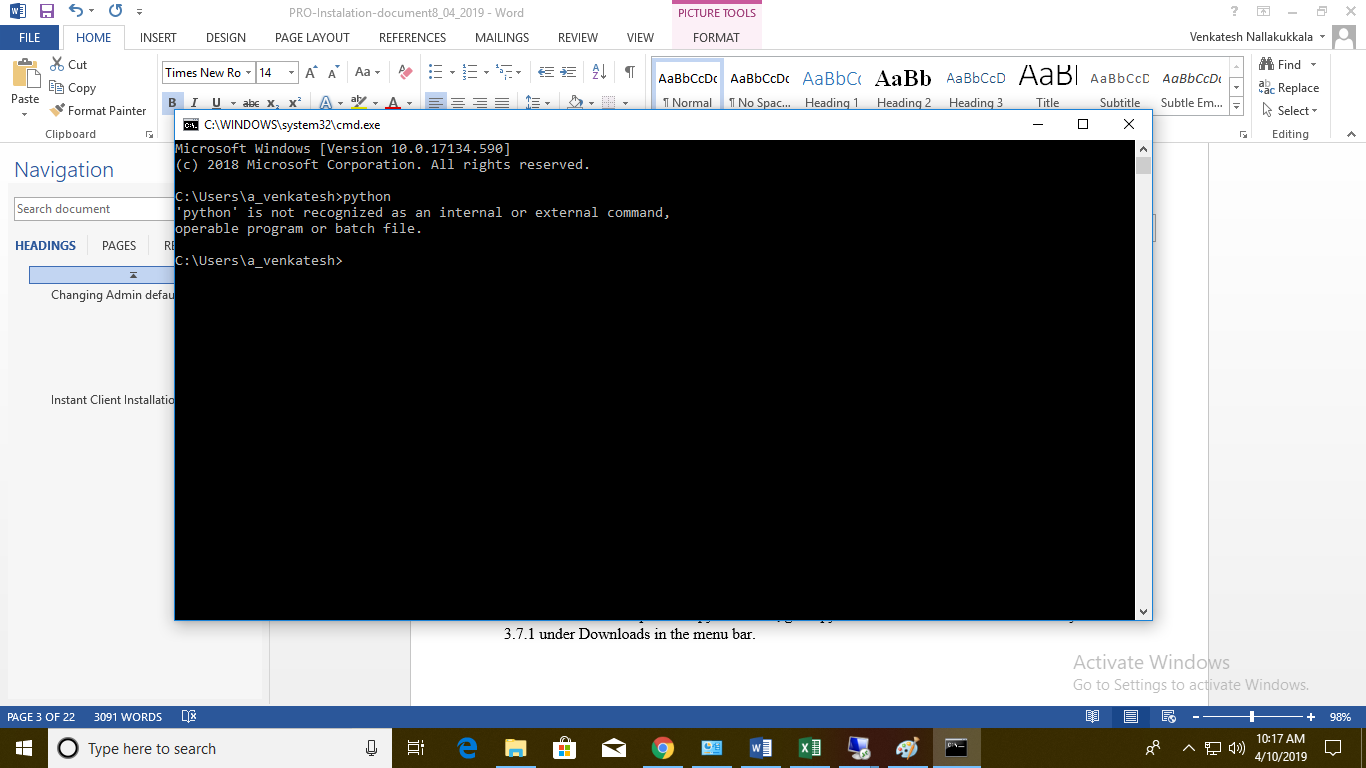
The objective of this tutorial is to help you set up python on windows OS. If you want to install python 3.7.1 instead, you might want to check this other tutorial Python 3.7.1 download and install for windows.

**Step1.**

At the time of this writing, the latest stable version of python is, released on October 20th, 2018. This tutorial will assist you in the download and install of python on windows. It is for both 32-bit and 64-bit versions of python.

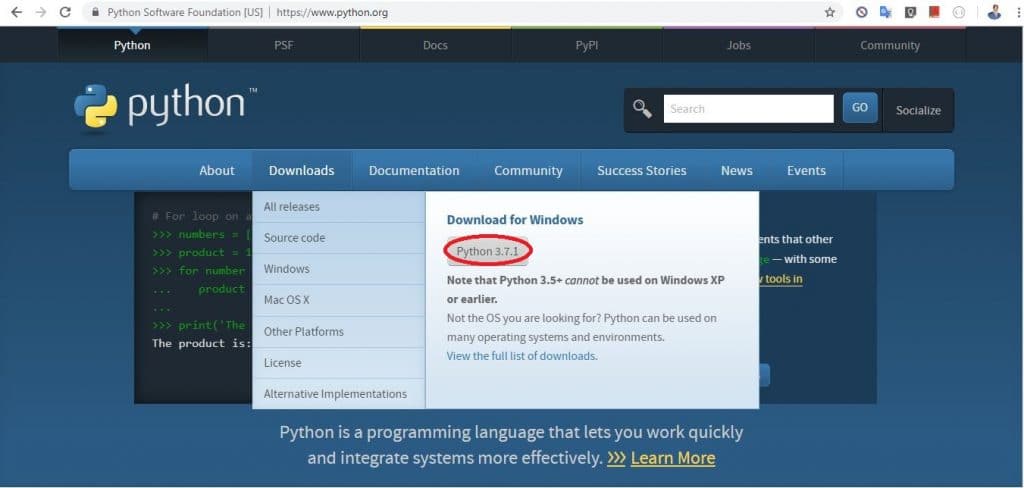
This installation process will automatically install IDLE, pip, and documentation as well and will also create shortcuts and file associations so that you don’t have to set up environment variables after the completion of installation.

Before starting, check if python is already installed on your computer. To do so, open command prompt and type “python” in it. If python is not already installed on your machine, you will see something like:

****

**Step2:**

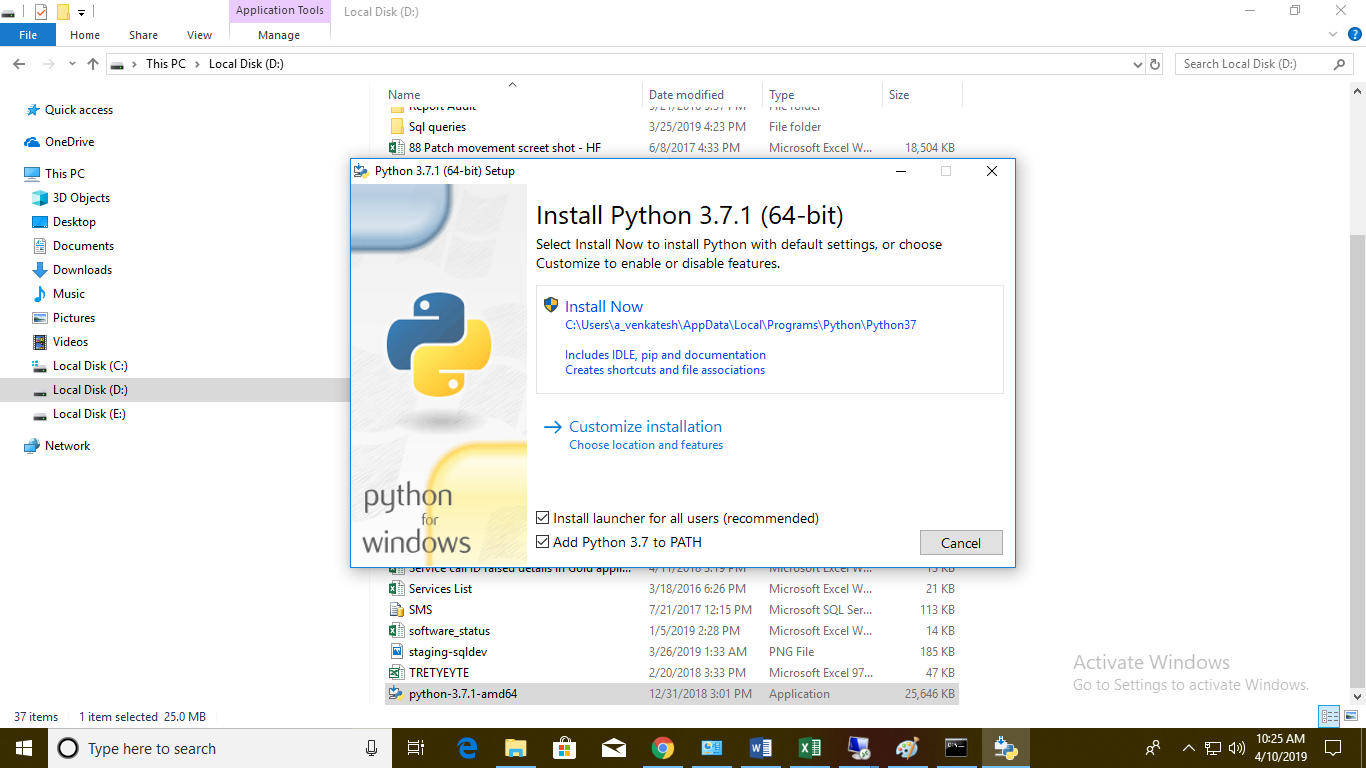
To download the setup file for python 3.7.1, go to python’s official website and click on Python 3.7.1 under Downloads in the menu bar.



Or, you can directly click on the link below to download the setup file.

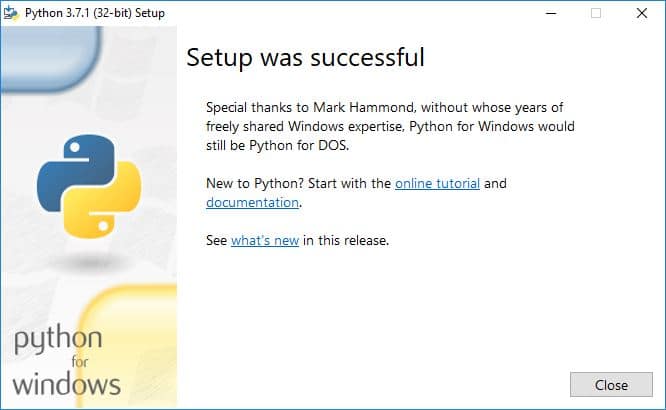
[**https://www.python.org/ftp/python/python.exe**](https://www.python.org/ftp/python/python.exe)

Once downloaded, locate the setup file under the name python.exe in the download folder and run it. You will see something like:



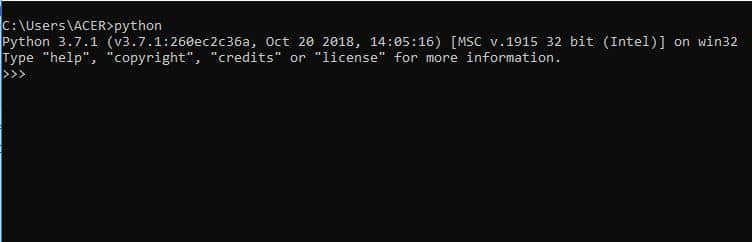
**Step3:**

By default, the Add Python to PATH option is unchecked, make sure it is checked then click on Install Now. If the setup is successful, you should see a window as below:



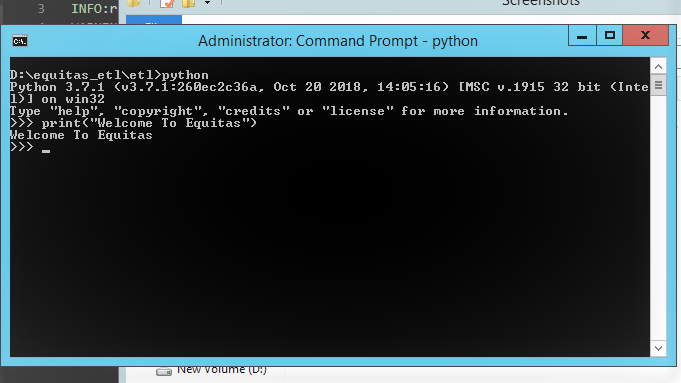
**Step4:**

Let’s check if python is successfully installed now. Open the command prompt and type “python” on it. If you haven’t closed the command prompt from earlier, you will need to close and reopen it. You will see something like:



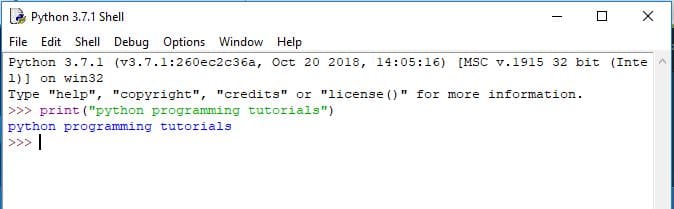
**Step5:**

Check that the python interpreter is working properly through command prompt:

****

**Step6:**

You can also search for IDLE and run python commands via.



**2. Visual C++ build tools**

**2.1 Purpose**

It provides an additional set of options that native C++ developers can install on top of the core MS Build components.

**Installation steps:**

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**3. Visual Studio Code +**

**3.1 Purpose**

Visual Studio Code is a source code editor developed by Microsoft for Windows, Linux and MACOS. It includes support for debugging, embedded Git control, syntax highlighting, intelligent code completion, snippets, and code refactoring.

**Installation steps:**

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**4.3 Magic Python:**

**4.3.1 Purpose**

This module uses ctypes to access the libmagic file type identification library. It makes use of the local magic database and supports both textual and MIME-type output.

**URL LINK:**

<https://files.pythonhosted.org/packages/42/a1/76d30c79992e3750dac6790ce16f056f870d368ba142f83f75f694d93001/python_magic-0.4.15-py2.py3-none-any.whl>

**Installation steps:**

1. First you need to download the software with extension of **.whl** file as we mentioned in above link.
2. If you downloaded **.whl(wheel)** extension file
3. You can directly run the below command for installing the software
4. **pip install filename.whl**

If you’re going to install .tar file first we need to install 7ZIP extractor to extract the software as normal.

**5. Dbeaver**

**5.1 Purpose**

DBeaver is a workbench for building SQL queries, editing and transferring data, viewing trees of objects, completing database administration tasks, monitoring database connection sessions, and a lot more. Having usability as its main goal, DBeaver offers: Carefully designed and implemented User Interface.

URL: <https://dbeaver.io/files/5.2.5/dbeaver-ce-5.2.5-win32.win32.x86_64.zip>

Step1: Download zipped Dbeaver software.

Step2: After downloading the software you have to extract that file as normal file.

Step3: Run the installer. Double-click the dbeaver-<version>.exe file that you downloaded to install DBeaver using the installer's default values.

6. SQLAlchmey 1.2.14

6.1 Purpose

SQLAlchemy is the Python SQL toolkit and Object Relational Mapper that gives application developers the full power and flexibility of SQL. SQLAlchemy provides a full suite of well-known enterprise-level persistence patterns, designed for efficient and high-performing database access, adapted into a simple and Pythonic domain language.

**Installation steps:**

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**7. Psycopg2 (version 2.7.6.1)**

**7.1 Purpose**

Psycopg2 is the most popular PostgreSQL database adapter for the Python programming language. Its main features are the complete implementation of the Python DB API 2.0 specification and the thread safety (several threads can share the same connection). It was designed for heavily multi-threaded applications that create and destroy lots of cursors and make a large number of concurrent INSERTs or UPDATEs.

URL: <https://pypi.org/project/psycopg2/>

The current psycopg2 implementation supports:

1. Python version 2.7
2. Python 3 versions from 3.4 to 3.7
3. PostgreSQL server versions from 7.4 to 10
4. PostgreSQL client library version from 9.1

**Installation steps:**

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**8. Future0.17.1**

**8.1 Purpose**

It allows you to use a single, clean Python 3.x-compatible codebase to support both

Python 2 and Python 3 with minimal overhead.

**Installation steps:**

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**9. Python executable installation and creating the executable file:**

PyInstaller is a normal Python package. You can download the archive from PyPi, <https://pypi.org/project/PyInstaller/>

**Step1**: Download the pyinstaller ==3.5.dev0+19d8a3789 by using the above link.

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

**step2: Install all the below dependency packages first. Then run Pyinstaller ==3.5**

1. Pywin32.ctypes == 0.2.0
2. Pywin32 ==224
3. Macholib ==1.11
4. Altgraph ==0.16.1
5. Pefile =2017.8.1
6. pypiwin32-223-py3-none-any.whl

**Step3**: After installing the all the above mentioned software’s follow the below steps to create the executable file

1. First you need go with specific path of your python file location then run the below command.

**Pyinstaller - - OneFile –w liability.py (name of the file).**

**Here:** This will create a standalone executable in the dist directory of your script folder. Don’t worry, if the folder doesn’t exist it will create one automatically.

Notice that we passed an argument “— OneFile”. This argument tells PyInstaller to create only one file. If you don’t specify this, the libraries will be distributed as separate file along with the executable.

**“-w”** w, --windowed, --no console

Windows and Mac OS X: do not provide a console window for standard i/o. On Mac OS X this also triggers building an OS X .app bundle. This option is ignored in \*NIX systems.

1. It will create **dist** , **build** directories and one **spec** file
2. In side dist directory the executable file will be create.
3. Before going to execute the executable file you have to setup **Spec file config.**
4. In spec file we have to add data’s file only remaining all automatically will come.

Datas= [('D:\\executable\\liability\\sql\\\*.sql', ‘sql’)]

**Here:**

* 1. 'D:\\executable\\liability\\sql\\\*.sql', it specifies the sql file location.
  2. Last parameter specifies the file to copy the all queries into sql file folder.

1. After editing the spec file we have to run the spec file

**D:\executable\etl>pyinstaller liability. Spec**

1. After completion of the all the above steps then run the below command from the command line where the file got placed.

**D:\executable\etl\dist>liability.exe**

1. Background job will run the process of your task.
2. Then go and check particular Database which database are using.
3. You have one more option to check the status of the task through the **Log file**.

**Note:** FileNotFoundError: [Errno 2] No such file or directory: 'C:\\Users\\etladmin\\AppData\\Local\\Programs\\Python\\Python37\\lib\\site-packages\\sqlalchemy-1.2.18-py3.7-win-amd64.egg\\sqlalchemy\\sql\\ddl.py'

Usually you will face this issue while generating the executable file. The issue is files reading by default it will come up with “.” Symbol.

We have to change that in to “-“ symbol because the sqlalchemy software extension is “**sqlalchemy-1.2.18-py3.7-win-amd64.egg”**

**Solution :**

Change the period symbol into “-“ in line number is 60.Below path is file location path to change.

**C:\Users\etladmin\AppData\Local\Programs\Python\Python37\Lib\site-packages\PyInstaller-3.4-py3.7.egg\PyInstaller\hooks\hook-sqlalchemy.py**

**10. Rundeck Installation:**

**Install on windows:**

The following guide will help you to install Rundeck PRO or Open Source as a Service on a Windows Server. It uses the Rundeck executable war installer.

Before going to installing the Rundeck software we have to install java because Rundeck is running in the java platform.

**Step1:** First download the java SE software into local and set the java path where your java file is installed .Follow the below steps to install and set the path.

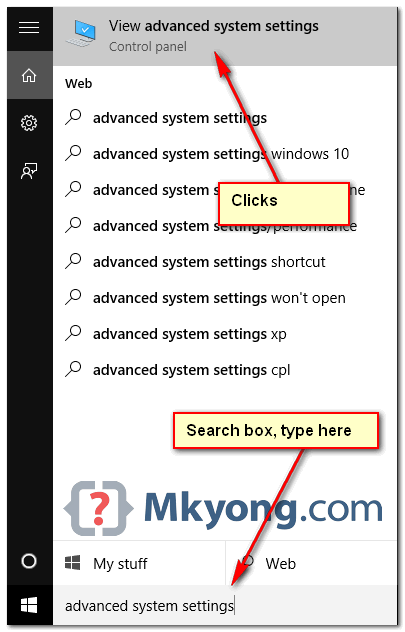
**Step 2:** Go to the below link choose the software windows and accept the license agreement then download the java .exe or zip file .If you download the zip file make sure before going to install extract that file .

<https://www.oracle.com/technetwork/java/javase/downloads/jdk11-downloads-5066655.html>

**Step 3:** then follow the below steps to setup the path.

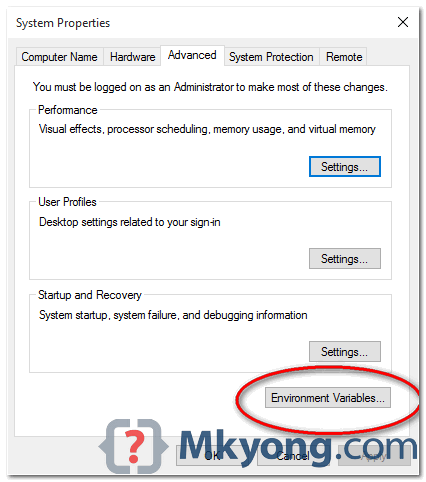
**1. Advanced System Settings**

Type advanced system settings in the search box (beside the Windows start button), clicks View advanced system settings.



**2. Environment Variables**

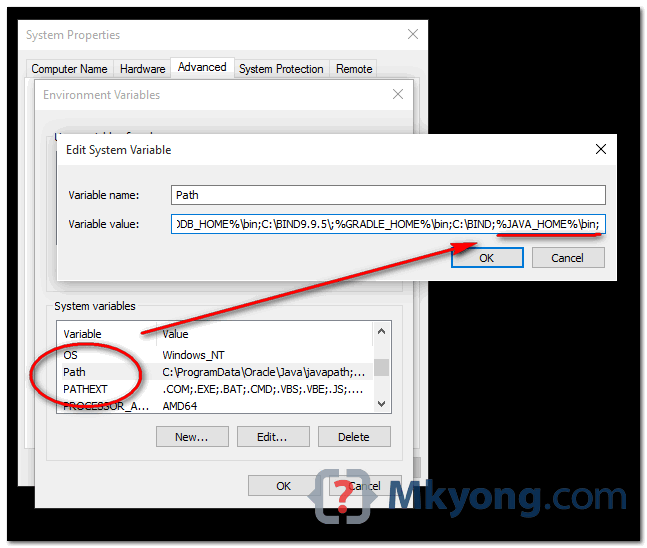
Select Advance tab, clicks Environment Variables



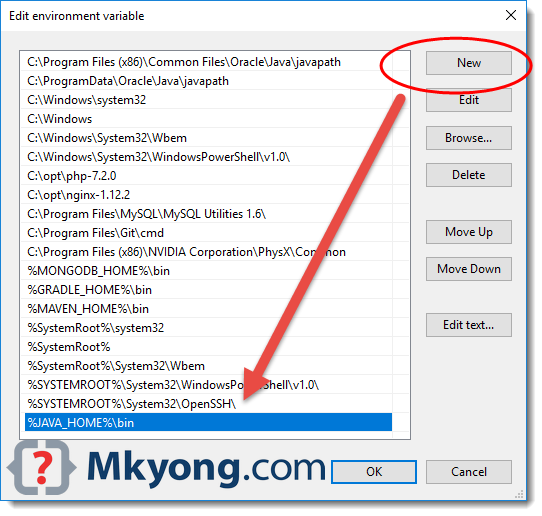
**3. Update PATH**

In System variables, find PATH, clicks edit... button:

* 1. In old version of Windows, it will prompt you below dialog box to edit the values directly, append this %JAVA\_HOME%\bin; to the end of the line.



4.2 In latest Windows 10, it will prompt you below dialog box, clicks on **new button**, and add this %JAVA\_HOME%\bin



**Note:** Puts the %JAVA\_HOME%\bin in PATH make all the Java’s commands (java, java, jstack and etc) are accessible from everywhere.

5. Test

**Open a command prompt, type:**

C:\rundeck>java --version

Java 11.0.1 2018-10-16 LTS

Java(TM) SE Runtime Environment 18.9 (build 11.0.1+13-LTS)

Java Hotspot(TM) 64-Bit Server VM 18.9 (build 11.0.1+13-LTS, mixed mode)

**Step 4:** Once you complete installation and path setup. Then follow the below steps.

4.1. Download the latest rundeck-3.0.9-20181127.war software from the below link.

<https://rundeck.org/downloads.html>

4.2. Choose a root directory (e.g. C :). This will be your %RDECK\_BASE% (ex: C:\rundeck)

4.3. Place Rundeck-[edition]-X.X.X.war in that directory

4.4. Go to the %RDECK\_BASE% (e.g. C :) folder and launch the installationof Rundeck.

1. set RDECK\_BASE=C:\rundeck
2. cd %RDECK\_BASE%
3. java -jar rundeckpro-[edition]-X.X.X.war

4.5. Windows launcher

After Rundeck started, stop the process.

* 1. Create a bat file (e.g. start\_rundeck.bat) and place it under %RDECK\_BASE%

4.7. Run the below commands from the CMD Prompt.

1. Set CURDIR=C: \Rundeck (based on your current dir. path.)

2. After successfully ran the above command then add the these commands in start\_rundeck.bat

Set CURDIR=%~dp0

Call %CURDIR%etc\profile.bat

java %RDECK\_CLI\_OPTS% %RDECK\_SSL\_OPTS% -jar rundeck-3.0.9-20181127.war --skipinstall -d >> %CURDIR%\var\logs\service.log 2>&1

* 1. Download nssm.exe

URL: <https://nssm.cc/release/nssm-2.24.zip>

4.9. Extract the zip file and copy the nssm.exe file

4.10. Place the executable under %RDECK\_BASE% (you can place it elsewhere, but for the sake of the example let’s use always the root dir.)

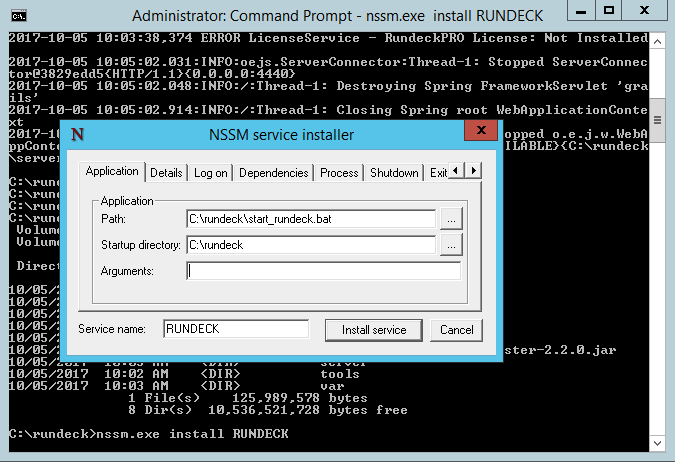
**Step5:** Open a prompt and issue these commands (Administrator mode required to install a service)

Cd C:\rundeck

Nssm.exe install RUNDECK

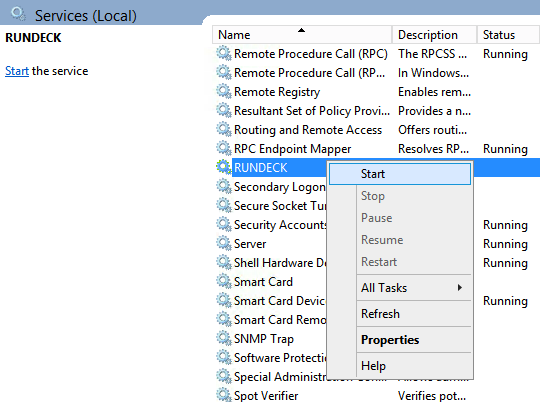
**Step6:**

The GUI pops up, set “path” as %RDECK\_BASE%\start\_rundeck.bat, startup directory as %RDECK\_BASE% (optionally set “low” on the process tab, under priority, to avoid server cpu spike when starting Rundeck)

NSSM Installer

**Step7:**

Go to the service management console (services. MSc) and you’ll find RUNDECK listed as a service. Starting it will start the Rundeck process.



**Step8:**

**Service Management Console**

Set JVM memory heap on %RDECK\_BASE%\etc\profile. Replace the RDECK\_CLI\_OPTS variable with the amount of memory that you need, for example:

....

Set RDECK\_CLI\_OPTS=-Xms512m -Xmx2048m

Set RD\_LIBDIR=%RDECK\_BASE%\tools\lib

**Step9:**

**Login to the GUI**

By default, Rundeck will be installed in port 4440. To access Rundeck, go to the following URL: ttp: //servername:4440.

Normally, it is necessary to modify the grails.serverURL property in %RDECK\_BASE%\server\config\rundeck-config.properties to access the GUI. Change grails.serverURL to the server name or IP (172.18.30.22).

9.1 After updating the **rundeck-config.properties** you will need to restart the Rundeck service for the change to take effect. (In the command prompt run this command **C:\rundeck>nssm restart Rundeck)**

* 1. After the restart, you can access the GUI at [**http://eqetlappuat01:4440/**](http://eqetlappuat01:4440/)
  2. Open this URL in internet explorer in the App server[**http://eqetlappuat01:4440/**](http://eqetlappuat01:4440/)

**Step10:**

**Login page**

The default username and password is **“admin”**

## Changing Admin default Password

In the case if you would like to change the default admin password, it’s so easy, edit the file **C:\Rundeck\server\config/realm.properties**,Change the second field ‘admin’ with a secure password as below:

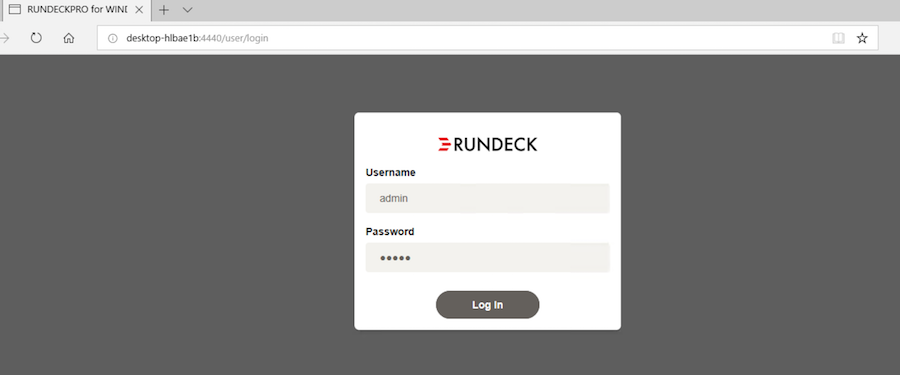
**C:\Rundeck\server\config/realm.properties**

# This sets the default user accounts for the Rundeck app

Admin: **YOUR\_NEW\_PASSWORD**, user, admin, architect, deploy, build

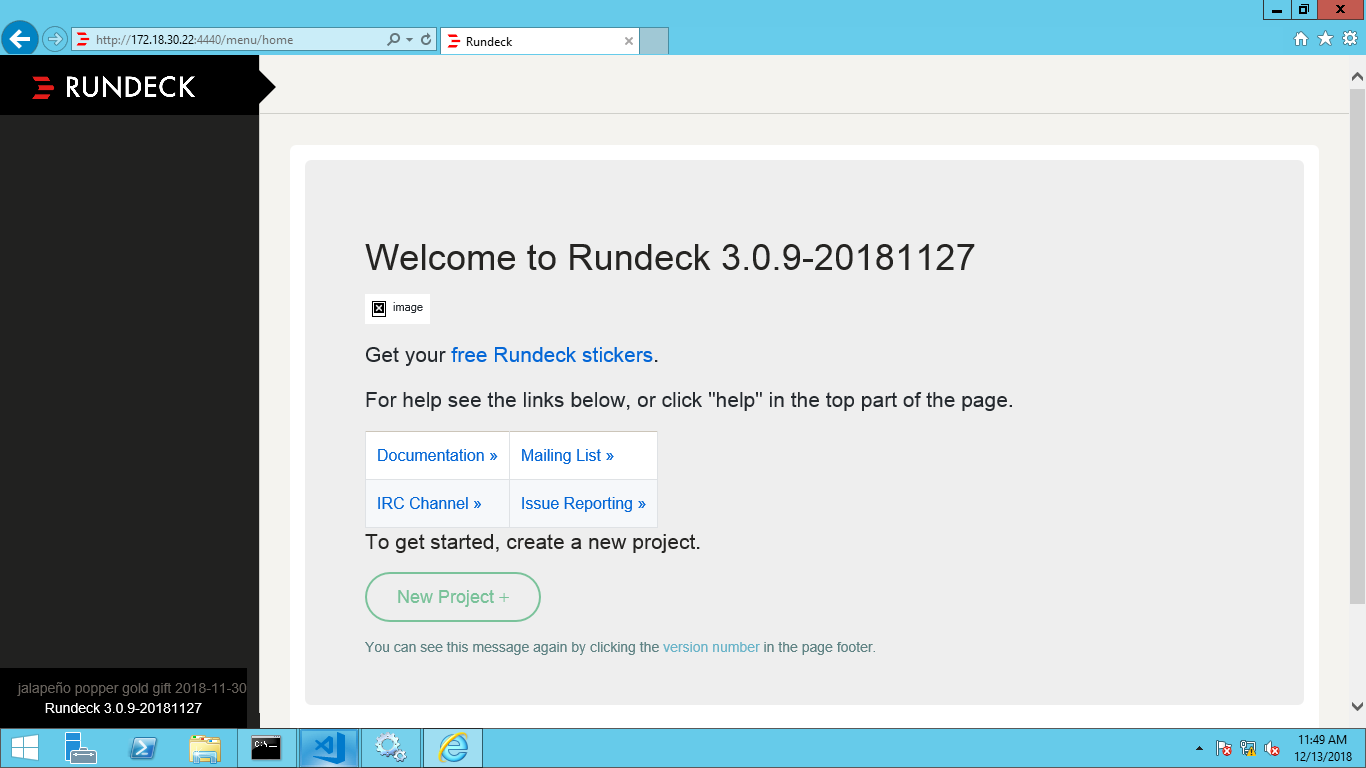
– After changing the default password, you have to restart the Rundeck:

**C:\Rundeck\nssm restart Rundeck**



**Step11:**

After login successfully you will get the below attached web page.



**Start-Stop instances**

Use the Service Windows GUI to start/stop/restart the instance OR

From the command prompt nssm <service name> Rundeck

**Hare:** service names are Start/stop/restart

11. Oracle Remote Database Connection through Python

Pre- requirements

11.1 instantclient-basic-windows.x64-12.1.0.2.0

11.2 instantclient-odbc-windows.x64-12.1.0.2.0

**Follow the below steps to config oracle**

<https://www.oracle.com/technetwork/topics/winx64soft-089540.html>

**Oracle Instant Client ODBC Installation Notes:**

Oracle's Instant Client ODBC software is a standalone package that offers the full functionality of the Oracle ODBC driver (except the Oracle service for Microsoft Transaction Server) with a simple install.

**Installing Oracle Instant Client Basic:**

Instant Client ODBC requires the Oracle Instant Client Basic or Basic Light package (depending on your locale requirements) also be installed. Download the desired package from OTN for your operating system and follow the installation instructions on the download page. For example, unzip the package to C:\Oracle\instantclient\_12\_1 on Windows

**Install Oracle Client**

Using cx\_Oracle requires Oracle Client libraries to be installed. These provide the necessary network connectivity allowing cx\_Oracle to access an Oracle Database instance. Oracle Client versions 18, 12 and 11.2 are supported.

If your database is on a remote computer, then download the free Oracle Instant Client “Basic” or “Basic Light” package for your operating system architecture.

Alternatively use the client libraries already available in a locally installed database such as the free Oracle XE release.

Oracle Instant Client Zip Files

To use cx\_Oracle with Oracle Instant Client zip files:

1. Download an Oracle 18, 12, or 11.2 “Basic” or “Basic Light” zip file: 64-bit or 32-bit, matching your Python architecture.
2. Unzip the package into a directory that is accessible to your application. For example unzip instantclient-basic-windows.x64-18.3.0.0.0dbru.zip to C:\oracle\instantclient\_12\_1.
3. Add this directory to the PATH environment variable. For example, on Windows 7, update PATH in Control Panel -> System -> Advanced System Settings -> Advanced -> Environment Variables -> System Variables -> PATH. The Instant Client directory must occur in PATH before any other Oracle directories.
4. Restart any open command prompt windows.
5. To avoid interfering with existing tools that require other Oracle Client versions, instead of updating the system-wide PATH variable, you may prefer to write a batch file that sets PATH, for example:
6. REM mypy.bat

SET PATH=C:\oracle\instantclient\_12\_1;%PATH%

Python %\*

1. Invoke this batch file every time you want to run python. Alternatively use SET to change your PATH in each command prompt window before you run python.
2. We have to check whether **VS 2010** is installed or not before going to setup the oracle ODBC.
3. Oracle Instant Client libraries require a Visual Studio redistributable with a 64-bit or 32-bit architecture to match Instant Client’s architecture. Each Instant Client version requires a different redistributable version:

**For Instant Client 12.1 install VS 2010**

1. If you intend to co-locate optional Oracle configuration files such as tnsnames.ora , sqlnet.ora or oraaccess.xml with Instant Client, then create a network\admin subdirectory, for example C:\oracle\instantclient\_12\_1\network\admin.
2. This is the default Oracle configuration directory for executables linked with this Instant Client.
3. Alternatively, Oracle configuration files can be put in another, accessible directory. Then set the environment variable TNS\_ADMIN to that directory name.

**The procedure to install Instant Client ODBC on Windows is:**

1. Download the Instant Client ODBC package. Unzip it in the same directory as your Basic or Basic Light package.
2. Execute odbc\_install.exe with the Administrator privilege from the Instant Client directory. If Instant Client is 11g or lower, start the command prompt with the Administrator privilege.

**The procedure to uninstall Instant Client ODBC on Windows is:**

1. Remove the DSN associated with the Oracle ODBC driver in the ODBC Data Source Administrator (odbcad32) console.
2. Execute odbc\_uninstall.exe from the Instant Client ODBC directory.
3. Delete all files and directories in the Instant Client ODBC directory.

**12. Instantclient-sqlplus-windows.x64-12.1.0.2.0**

The sqlplus instance client is used for connecting the oracle database from the command prompt.

1. Download an “Instantclient-sqlplus” zip file: 64-bit or 32-bit, matching your Python architecture.
2. Unzip the package into a directory that is accessible to your application. For example unzip instantclient-basic-windows.x64-18.3.0.0.0dbru.zip to C:\oracle\instantclient\_12\_1.
3. Open the command prompt with admin privileges the run the below command

**Sqlplus username/password@hostname:port-number/service name**

1. Then press enter. We can get the Sql> /command prompt.

**13. Pandas Installation**

Pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

Pandas is a NumFOCUS sponsored project. This will help ensure the success of development of pandas as a world-class open-source project, and makes it possible to donate to the project.

These are the dependency packages for the panda’s dataframe in windows.

13.1 Six

13.2 Python –dateutil

13.3 Pytz

13.4 Psutil

Follow the same process for all the above dependency packages.

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**

14. Numpy

1. First you need to download the software with extension of either .tar file or .whl file
2. If you downloaded **.tar** extension file follow the below steps.
3. After downloading the software you have to extract the software by using 7Zip.
4. Then go to setup.py file container directory then run the software on command line by using the below command.
5. **Python setup.py build**
6. After successfully build. Run the below command
7. **Python setup.py install**
8. If you downloaded **.whl(wheel)** extension file
9. You can directly run the below command for installing the software
10. **pip install filename.whl**